

公共数据开放对企业绿色创新的统计研究

刘晓春

广东财经大学 经济学院, 广州, 510320;

摘要: 本文以 2003-2022 年我国 287 个地级市相关数据为基础, 采用多期双重差分模型考察各地区公共数据开放前后, 企业在绿色创新的差异性表现。研究发现, 公共数据开放对企业绿色创新存在正向推动的作用, 尤其在策略性绿色创新的表现更活跃, 这得益于人力资本的提升、创业氛围的增强和环境规制的倒逼作用。

关键词: 公共数据开放; 绿色创新; 数据要素

DOI: 10.69979/3029-2700.25.03.024

引言

近年来, 我国经济已由高速增长阶段转向高质量发展阶段, 加快推动经济社会发展全面绿色转型已经形成高度共识。2021 年 3 月, “十四五”规划基于“双碳”战略目标多次提出绿色转型概念, 实施绿色技术创新攻关行动, 推进重点行业和重要领域的绿色化转型。在此背景下, 绿色创新不仅是我国实现高质量发展的内在要求, 也是“双碳”战略目标的关键驱动力。作为新时代的重要生产资源, 数据要素的高效率与低耗能特性, 不仅有助于减少能源消耗和碳排放, 减轻环境压力, 更能为推动绿色创新提供有力支持^[1]。公共数据作为数据要素的重要组成部分, 通过为企业和科研机构提供丰富的数据资源从而有助于企业绿色创新的发展。鉴于此, 政府开放的数据作为数据要素存在一定的价值, 符合推动绿色创新的需要, 能为企业进行绿色创新提供新视角。

当前, 学术界对公共数据开放与企业绿色创新之间的研究已有一定讨论。国内外学者从不同维度深入探讨了公共数据开放对绿色创新溢出效应, 普遍认为, 公共数据开放能够有效促进企业加大在环境保护和创新能力提升方面的投入和重视。但是, 已有研究主要聚焦于绿色创新总体, 验证了公共数据开放为企业绿色创新水平提升带来了更多机会, 也有文献在公共数据开放对两类绿色创新的差异化作用效果进行了探讨, 发现相对于策略性绿色创新, 公共数据开放对实质性绿色创新的促进作用表现得更为明显^[2], 但未进行进一步分析。

综上所述, 尽管已有一些研究关注公共数据开放对两类绿色创新的差异化影响, 但整体上研究数量仍显不足, 有必要进一步探究这一命题。具体而言, 本文从“人力资本-创业氛围-环境规制”三重维度出发, 深入剖析

公共数据开放推动企业绿色创新发展的内在逻辑和实现途径。

1 理论分析与理论假说

1.1 公共数据开放对企业绿色创新的直接影响

随着数据开放平台的兴起, 公共数据作为一种新型生产要素, 为企业开辟了新的发展契机。而在推进绿色创新的过程中, 数据开放通过可访问性释放数据价值, 带动企业创新能力的提升。如公共数据开放创造更公平的竞争环境并促使企业参与绿色创新。因此, 本文提出研究假说 H1。

H1: 公共数据开放后能正向推动企业绿色创新的发展。

1.2 公共数据开放对两类绿色创新的差异化影响

绿色创新根据不同的动机可以分为“实质性绿色创新”和“策略性绿色创新”由于市场竞争较为剧烈, 企业更可能选择策略性绿色创新。但是, 因为公共数据开放平台提供的数据为企业提供丰富的数据资源, 因此企业也有可能会加大对于实质性绿色创新的投入。因此, 本文提出研究假说 H2。

H2: 相较于实质性绿色创新, 公共数据开放对策略性绿色创新的正向影响更为明显, 或至少相当。

1.3 公共数据开放对企业绿色创新的影响机制

公共数据开放推动企业绿色技术创新发展的过程中, 人力资本、创业活力和环境规制起到中介的作用。

1.3.1 增加人力资本

在公共数据开放引致良好市场环境的加持下, 高质量人才要素流入, 有助于知识与技术的交流融合^[3]。而

人才的聚集能够提升企业发展能力,进而推动企业绿色创新^[4]。因此,提出研究假说 H3a。

H3a: 公共数据开放通过增加人力资本进而推动绿色创新。

1.3.2 培养创业氛围

政府数据开放通过制度性交易成本降低和信息搜寻匹配效率优化来推进城市创业活力提升^[3]。创新创业水平活力的提升,有助于形成创业氛围,激发企业家精神,进而推动绿色技术创新。因此,提出研究假说 H3b。

H3b: 公共数据开放通过提升创业氛围进而推动绿色创新。

1.3.3 加强环境规制

公共数据开放能通过增强信息透明度,进而优化监管环境。这将有利于倒逼产业转型升级、加快高耗能高污染企业去产能、提升企业的绿色创新^[5]。因此,提出研究假说 H3c。

H3c: 公共数据开放通过加强环境规制进而推动绿色创新。

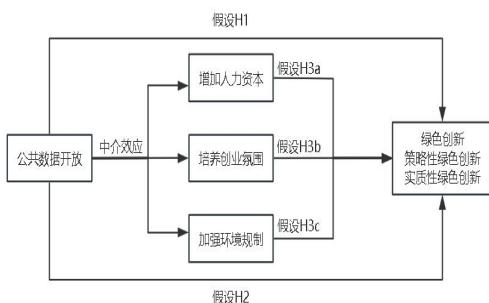


图 1 公共数据开放推动绿色创新研究机制框架

2 研究设计

2.1 模型设定

根据前文分析,公共数据开放对绿色创新可能存在正向影响,为验证实际的效果,构建如下面板模型:

$$Y_{it} = a_0 + a_1 DATA_{it} + a_2 CONTROLS_{it} + \mu_i + \delta_t + \varepsilon_{it} \quad (1)$$

式(1)中的被解释变量 Y_{it} 表示企业绿色创新水平,本文用实质性绿色创新(Invest)、策略性绿色创新(Practical)及其加总的绿色创新(Total)来表示。主要解释变量 $DATA_{it}$ 为公共数据开放平台上线前后的虚拟变量。为了获得更稳健的回归结果,本文加入了城市固定效应(μ_i)和时间固定效应(δ_t)。最后,随机

扰动项用 ε_{it} 表示。

2.2 变量定义

2.2.1 被解释变量

绿色创新(Y_{it})主要参考王家庭(2024)^[6]的做法,选取专利申请量而非获取量作为企业绿色创新的指标。

2.2.2 解释变量

公共数据开放($DATA_{it}$)的取值逻辑如下:如果某地级市,在 t 年已经上线了公共数据开放平台,则取值 1,表示存在政策影响;否则取值 0,表示不存在政策的影响。

2.2.3 控制变量

变量定义如下:政府科技支出强度(Govrd):以政府科学技术支出占政府财政支出的比重衡量;财政分权(Fd):依据地区财政支出与财政收入的比值计算;基础设施(Road):采用人均城市道路面积作为量化指标;外商直接投资(Fdi):通过地区实际利用外商直接投资额来表示;人力资本(Hum):普通高等学校在校学生数来表示;经济发展水平(Agdp):通过城市人均GDP反映。

2.3 数据来源与描述性统计

本文的数据来源主要包括:(1)地级市层面的绿色创新数据来源于国家知识产权局。(2)关于各地级市公共数据开放平台上线情况,参考《2024 中国地方公共数据开放利用报告》。(3)地级市层面的控制变量和机制变量数据主要来源于《中国城市统计年鉴》,对于部分缺失值,主要采用线性插值的方法进行补充。

变量数据的描述性统计情况如下:

表 1 变量的描述性统计分析

变量	样本数	均值	标准差	最小值	最大值
Total	5,740	4.286	15.455	0.010	297.540
Invest	5,740	2.056	8.813	0.010	215.400
Practical	5,740	2.231	7.078	0.010	109.590
Govrd	5,740	1.366	1.555	0.000	0.144
Fd	5,740	2.873	1.878	0.582	18.399
Road	5,740	15.872	8.193	0.390	215.660
Fdi	5,740	0.762	2.661	0.000	100.243
LnHum	5,740	10.319	1.495	2.341	14.086

LnAgdp 5,740 10.329 0.843 7.153 12.456

3 实证分析

3.1 公共数据开放对绿色创新影响的基准回归结果

可见公共数据开放对绿色创新存在正向驱动作用,验证了假说 H1。从公共数据开放对于实质性绿色创新的系数分别为 5.155 和 3.261, 对于策略性绿色创新的系数分别为 4.942 和 3.254, 假说 H2 得到验证。

表 2 基准回归

变量	(1)	(2)	(3)	(4)	(5)	(6)
	绿色创新	策略性绿色创新	实质性绿色创新	绿色创新	策略性绿色创新	实质性绿色创新
DID	10.097 *** (3.709)	4.942* ** (4.036)	5.155* ** (3.324)	6.515** * (3.514)	3.254** * (3.868)	3.261** * (3.079)
_cons	3.191* ** (10.819)	1.695* ** (12.77)	1.497* ** (8.907)	726.474 *** (4.394)	350.926 *** (5.653)	375.766 *** (3.449)
控制变量	否	否	否	控制	控制	控制
时间&城市	否	否	否	控制	控制	控制
N	5740	5740	5740	5740	5740	5740
R2	0.596	0.587	0.582	0.678	0.673	0.654

注: ***、**、*分别表示在 1%、5%、10% 的统计水平上显著, 括号中数值表示统计量 t 值。下文同。

3.2 平行趋势检验

由图 2 可知, 公共数据开放前, 处理组和对照组的绿色创新和两类绿色创新并不存在显著差异, 而在公共数据开放后几年, 处理组和对照组的绿色创新存在显著的差异, 即平行趋势检验通过, 可以进行多重差分方法进行检验。

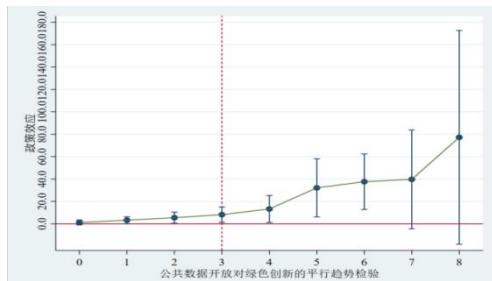


图 2 公共数据开放对绿色创新的平行趋势检验

3.4 公共数据开放对绿色创新影响的稳健性检验

3.4.1 剔除直辖市

剔除直辖市后结果如表 3 第 (1) 列所示, 由显著性水平未发生明显变化, 说明回归结果较为稳健。

3.4.2 更换主要解释变量

以绿色创新的授权量替代绿色创新的申请量进行回归, 表 3 的第 (2) 列汇报了相应结果。可以看出, 结论依旧显著, 验证了回归结果的稳健性。

3.4.3 克服离群值的影响

对各个变量的数据进行 1% 的缩尾处理, 回归结果如表 3 第 (3) 列所示。主要解释变量的系数为正且依旧显著, 说明公共数据开放显著正向影响对绿色创新。

表 3 公共数据开放对企业绿色创新的稳健性检验

变量	剔除直辖市	更换主要解释变量	克服离群值的影响
	绿色创新		
DID	4.569*** (3.333)	4.088*** (3.893)	4.088*** (3.893)
_cons	517.350*** (5.659)	469.253*** (4.503)	469.253*** (4.503)
控制变量	控制	控制	控制
时间&城市	控制	控制	控制
N	5660	5740	5740
R2	0.668	0.642	0.642

4 进一步分析

4.1 作用机制分析

下面进行中介效应分析, 验证 H2a、H2b 和 H2c 是否正确, 并构建以下模型:

$$M_{it} = \beta_0 + \beta_1 DATA_{it} + \beta_2 CONTROL_{it} + \mu + \delta + \tau_i \quad (2)$$

$$Y_{it} = \gamma_0 + \gamma_1 M_{it} + \gamma_2 CONTROL_{it} + \gamma_3 DATA_{it} + \mu + \delta + \nu_i \quad (3)$$

其中, M_{it} 表示机制变量, τ_i 和 ν_i 分别表示随机误差项。

4.1.1 增加人力资本

采用信息传输、计算机服务和软件业从业人员数与科研、技术服务和地质勘查业从业人员数之和占城市总从业人员的比重, 作为城市人力资本大小的反映。表 4 的第 (1) 列可知, 公共数据开放可以通过增加人力资本

本,推动绿色创新的发展。

4.1.2 培养创业氛围

将城市人口作为标准化基数,采用城市每百人口中的新注册企业数量衡量城市创业活力。由表4的第(2)列知公共数据开放能提升创业氛围,进而推动绿色创新发展。

4.1.3 加强环境规制

统计2003-2023年政府报告中与环境规则相关的词频,并除以文本总量,以此作为环境规制强度的衡量指标。由表4的第(3)列可知,公共数据开放通过加强环境规制水平,进而促进绿色创新的发展。

表4 公共数据开放与绿色创新的作用机制

变量	(1)	(2)	(3)
	人才资本	创业氛围	环境规制
DID	0.003** (2.096)	0.204* (1.956)	-0.002** (-2.000)
_cons	0.265** (2.034)	156.870 (1.397)	0.321 (1.570)
控制变量	控制	控制	控制
时间&城市	控制	控制	控制
N	4879	5453	5700
R2	0.815	0.253	0.147

5 研究启示

本文的研究结论对于发展公共数据开放平台和推动企业绿色创新有一定的启示。

第一,加快公共数据开放平台的建设,以发挥公共数据资源的价值。建设高质量的数据要素供给平台,借助数据的信息有效性、准确性和完整性,能较好地支持企业进行绿色创新。还应该充分认识到数据开放对绿色创新的正向作用,以最大化地发挥数据开放的效果,更好地推动绿色创新的发展。

第二,充分发挥“人力资本-创业氛围-环境规制”

三个方面的在公共数据推动企业绿色创新的中介作用。首先,充分发挥公共数据平台使得市场变得透明、增强市场监督力量和降低信息不对称的作用,营造公平竞争的环境和推动跨领域人才的聚集合作,吸引更多的人力资本,激发经济活力;其次,以市场为导向,通过与相关的行业进行合作,提供更多有价值可公开的数据资源,开拓创业者的市场机会和方向思路,营造充满创新活力的氛围;最后,充分发挥政府、公众和企业的监管作用,提高数据市场管理治理能力,倒逼企业完善绿色创新。

参考文献

- [1]于洋,陈放,王尔大.数据要素配置、新质生产力与区域绿色创新绩效[J].统计与决策,2024,40(17):5-11.
- [2]孟望生,范鼎浩,李丁.政府数据开放对城市绿色创新效率的影响——基于政府数据开放平台上线的准自然实验[J].技术经济,2024,43(09):1-17.
- [3]蔡运坤,周京奎,袁旺平.数据要素共享与城市创业活力——来自公共数据开放的经验证据[J].数量经济技术研究,2024,41(08):5-25.
- [4]申明浩,谭伟杰.数字化与企业绿色创新表现——基于增量与提质的双重效应识别[J].南方经济,2022,(09):118-138.
- [5]万晓乐,毕力文,邱鲁连.何种环境战略生态产生企业高绿色创新绩效?——基于SSE框架的组态分析[J].中国人口·资源与环境,2022,32(09):92-107.
- [6]王家庭,盛楠.数字化转型推动企业绿色创新——基于信息效应、资源效应与技术效应的分析[J].北京工业大学学报(社会科学版),2024,24(05):113-130.

作者简介:刘晓春(2000-),男,汉,湛江市,学生,硕士研究生,广东财经大学,社会经济统计

中图分类号: F124.3; X322 文献标识码: A